## SPECIFIC AIMS

This project proposes continued development of advanced interactive visualization and analysis software for experimental data, including and emphasizing data from state-of-the-art (<3Å resolution) cryo-electron microscopy (cryo-EM), with the ultimate goal of understanding how cells and their molecular machinery function. The enormous growth in recent years in both the size and complexity of biological data sets, and especially structural data at various scales in length and time has created new and significant challenges for biomedical researchers. New and innovative software tools are vital to the successful outcomes of the myriad NIH-funded experimental research projects. There is no indication that the growth in data size and complexity will abate, and thus approaches to integrating, interpreting, and otherwise making the best use of the data require continuing, focused attention. We propose to address this challenge via the following specific aims:

**Aim 1: Develop interactive visualization and analysis tools for atomic structures and associated data, including higher-order assemblies, conformational ensembles and trajectories, density maps, and sequences.**

Understanding the biological roles of molecular structures requires integrating a variety of associated data and analyses, with interactive visualization to link alternative views and diverse types of information with important structural elements. We will extend ChimeraX with a rich set of new features, for example: increased use of interactive 2D plots; a streamlined process for preparing structures for docking or further calculations by adding missing atoms and assigning charges; new sequence/structure tools, including modeling loops and generating sequence alignments from structure superpositions; and interactive localized molecular dynamics (MD) simulations to evaluate flexibility or the effects of mutations. To facilitate making compelling movies for publication and teaching, we will implement a timeline animation tool. Further, we will enhance the capabilities and performance of key tools, as well as adapt to the needs of the broader structural biology community as they arise.

**Aim 2: Develop interactive visualization and analysis tools for electron microscopy of molecular assemblies, cells and tissues.**

Profound advances in electron microscopy are extending it to atomic (1-2Å) resolution and to scales as large as the *Drosophila* brain, posing many new challenges in turning image data into biological insights. We will develop robust and widely useful analysis capabilities and make the latest innovative algorithms being developed in labs around the world accessible to the broad research community. We will provide interfaces to new machine-learning methods that capture heterogeneous states in single-particle cryoEM of molecular assemblies, and others that segment tomography of cells, and to tiled imaging of tissues that promises to reveal an order of magnitude more complexity in functional biological systems. These interactive tools will leverage new advances in visualization such as uniform manifold approximation and projection (UMAP) for seeing patterns in higher-dimensional data, from discerning flexible modes of molecular machines to characterizing cells from observed quantitative properties. Virtual-reality visualization and lighting of tomograms will help discern new structures. Significant effort will also go into making proven tools for 3D microscopy and models accessible to the widest audience possible.

**Aim 3: Provide a diverse software foundation enabling labs around the world to develop and distribute new molecular and cellular structure analysis software.**

Obstacles to new structure analysis methods getting beyond journal publication and into wide use are severe. New methods require a foundation of underlying software to read complex data, visualize it, provide intuitive user interfaces, and make it easily installed by researchers on the major computer operating systems. This aim will allow other labs to easily extend ChimeraX. Steps to achieve this include design and documentation of stable public software interfaces; creating programming tutorials on all aspects of reading data, computing, and adding user interfaces such as mouse modes, toolbars, panels and commands; and utilizing online software repositories for distribution and for developers to manage new releases and collect usage statistics, bug reports, and user feedback. Simplifying the development of production-quality implementations of new analysis methods will multiply the number of new ChimeraX tools available to researchers and advance the pace of scientific discovery.